SANDBOXAQ™

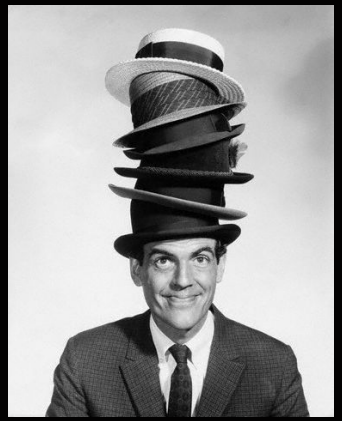
GOOD CHEMISTRY COMPANY

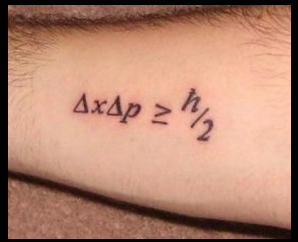
One million core simulation of PFAS chemistry

David Walker

1 About Me

A brief history…





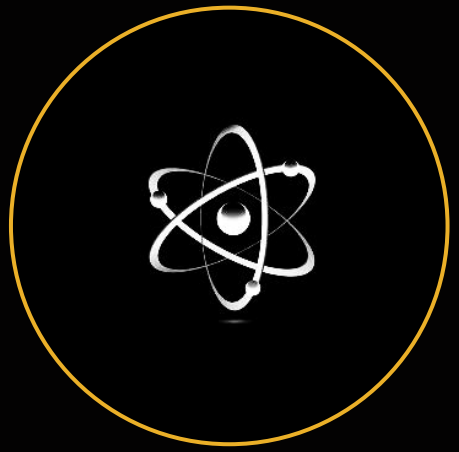
2

SandboxAQ

We solve large-scale enterprise challenges using the best of today's physics & AI platforms

Artificial Intelligence, World of bits

AI is powerful when there is enough training data

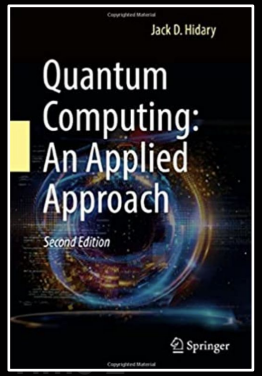


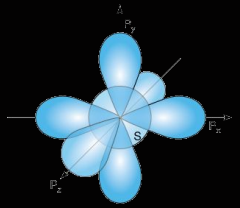
Quantum Physics, World of atoms

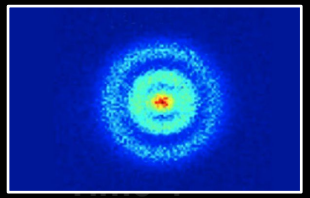
Physics tools create new data about the world around us

SandboxAQ Timeline

SANDBOXAQ™







2016
Sandbox founded inside Alphabet

2019
Seminal AI papers and quantum book published

2020
1st customers

Eric Schmidt joins

Quantum papers published

2021
Largest DFT computation to date - 100,000+ electrons

2022
Spin-out of Alphabet

Papers in Nature, PNAS, etc.

SANDBOXAQ™

We deliver market-ready AI & quantum solutions

No need to wait for Quantum Computers

Simulation & Optimization

Reduce the time and cost to discover and develop new molecules

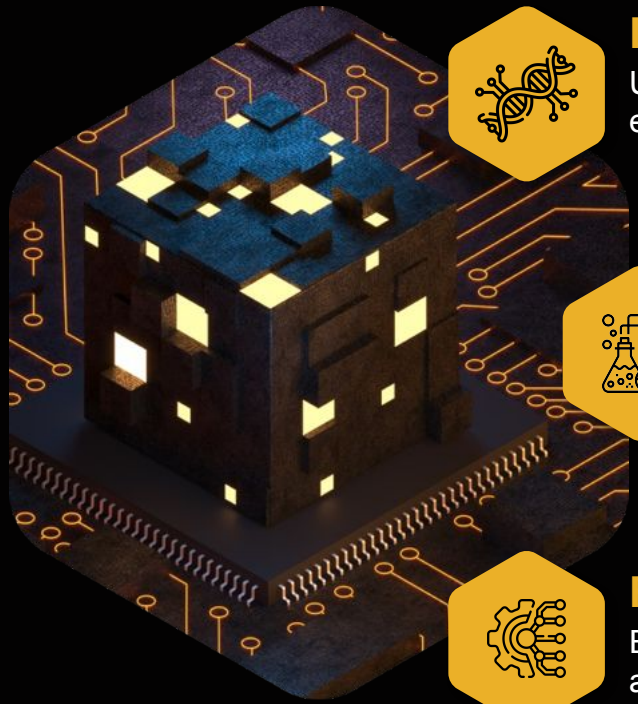
CyberSecurity & Communications

Protect sensitive IP & data against near- and long-term threats

Quantum Sensing

Generate new insights with AQ medical imaging hardware and software

Simulation & Optimization



Life Sciences

Unlocking complex biological systems and expediting novel therapeutic development

Chemicals & Materials

Advancing the discovery and design of innovative chemicals and materials

Finance & Logistics

Enabling unparalleled optimization, predictive analytics, and streamlined operations

3

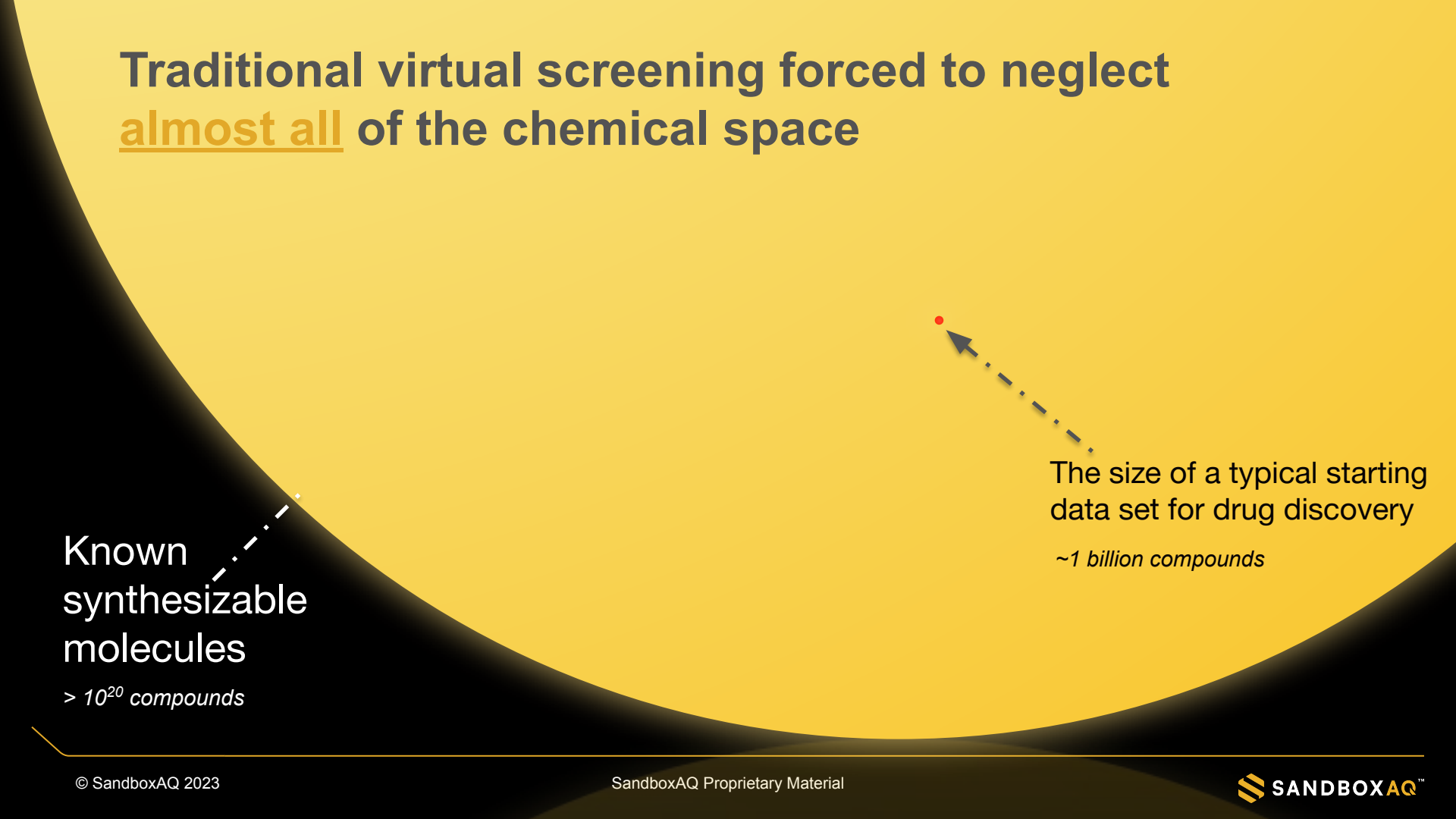
Good Chemistry

Our Mission

To make the world cleaner, healthier, and more sustainable by accelerating materials design

SANDBOXAQ™

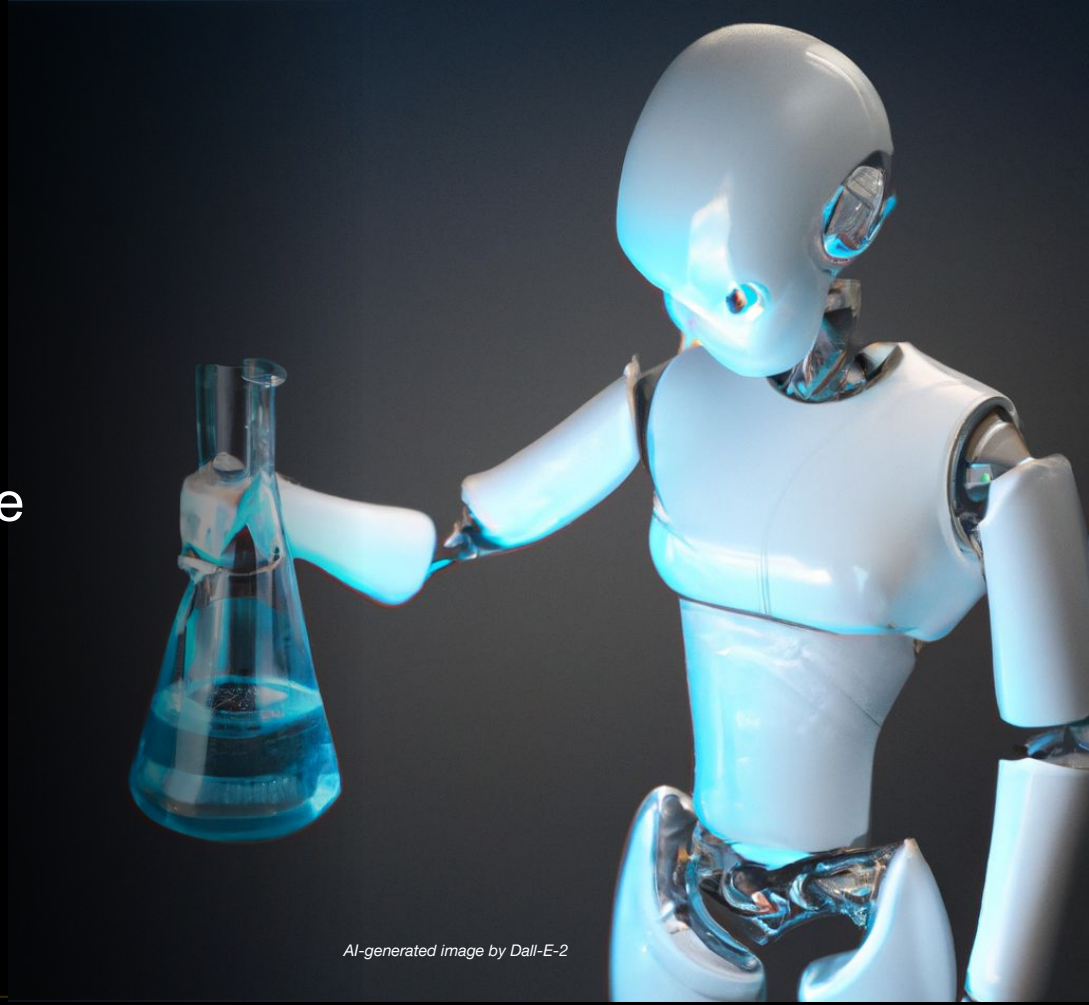
Traditional virtual screening forced to neglect almost all of the chemical space



GOOD CHEMISTRY
COMPANY

Making the intractable tractable in materials innovation using

Artificial Intelligence Trained on high-accuracy quantum chemistry data

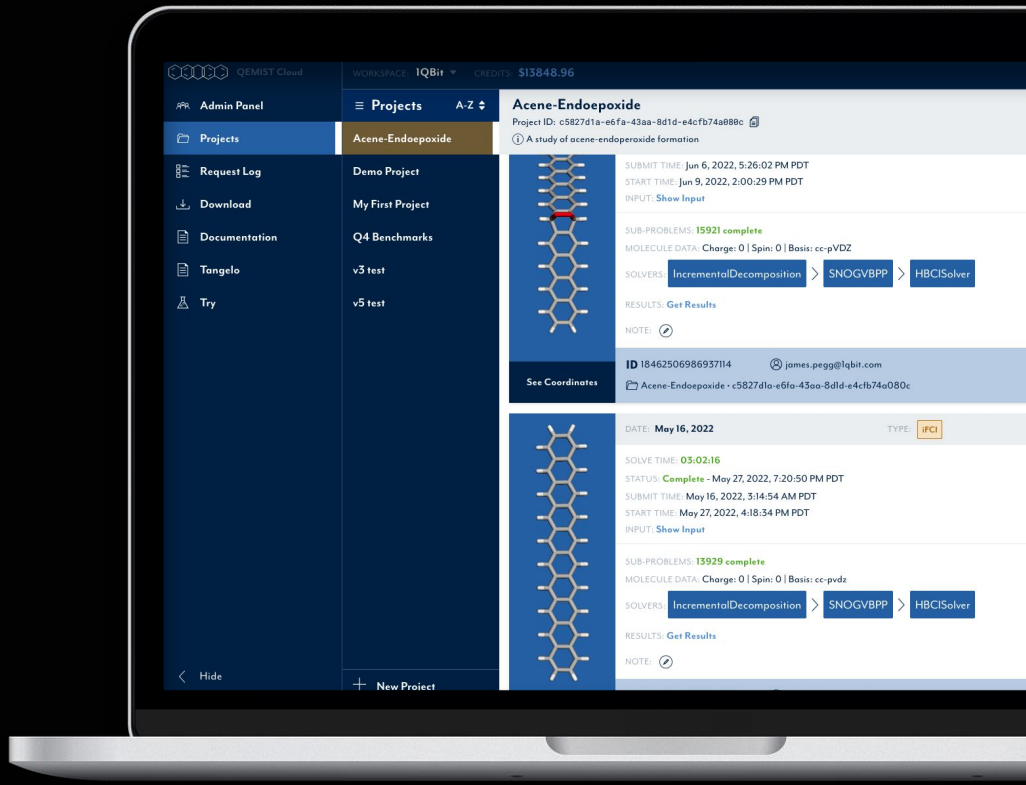


AI-generated image by Dall-E-2

SANDBOXAQ™

QEMIST Cloud

A collaborative virtual chemistry lab on the cloud for exploring the chemical space that was previously out of reach



aws PARTNER Qualified Software

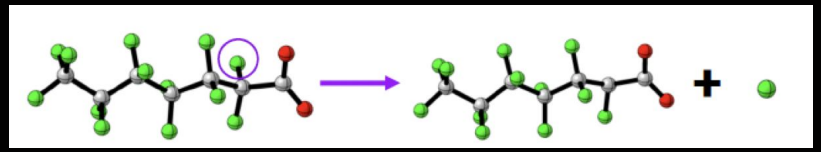
Qualified AWS Partner
QEMIST Cloud is AWS FTR reviewed

SANDBOXAQ™

4 The Task

One Million Core Simulation of PFAS Molecules

- PFAS (Per- and polyfluoroalkyl substances) are hazardous “forever chemicals”. they can bioaccumulate and cause adverse health (e.g. cancer) and environmental issues.
- Chemical degradation pathways for PFAS mainly involve breaking carbon-fluorine bonds.



Defluorination of PFOA molecule

- Identifying the best pathways, catalysts, and engineering solutions to the degradation/defluorination of PFAS compounds requires accurate simulations of PFAS molecules.

In collaboration with

aws

accenture

intel

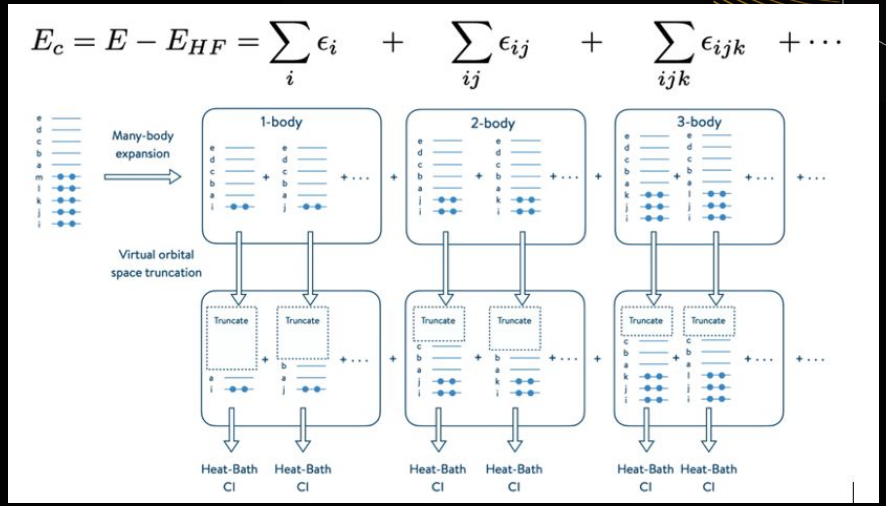
SANDBOXAQ™

An Intractable Problem

Exact solution requires the diagonalization of a 10^{151} x 10^{151} matrix

Computational Chemistry 101

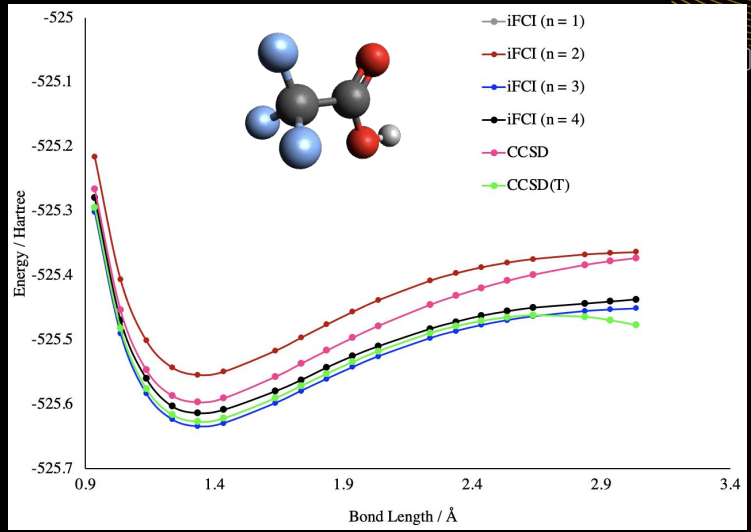
- Full Configuration Interaction (FCI) represents the formally exact solution to the electronic Schrödinger equation in a given basis set. However, It is intractable even for the smallest PFAS molecule, TFA (trifluoroacetic acid)
- The cloud native implementation of iFCI in QEMIST cloud offers a polynomial scaling approximation to FCI and provides the most accurate treatment of the PFAS bond breaking by taking advantage of cloud computing.



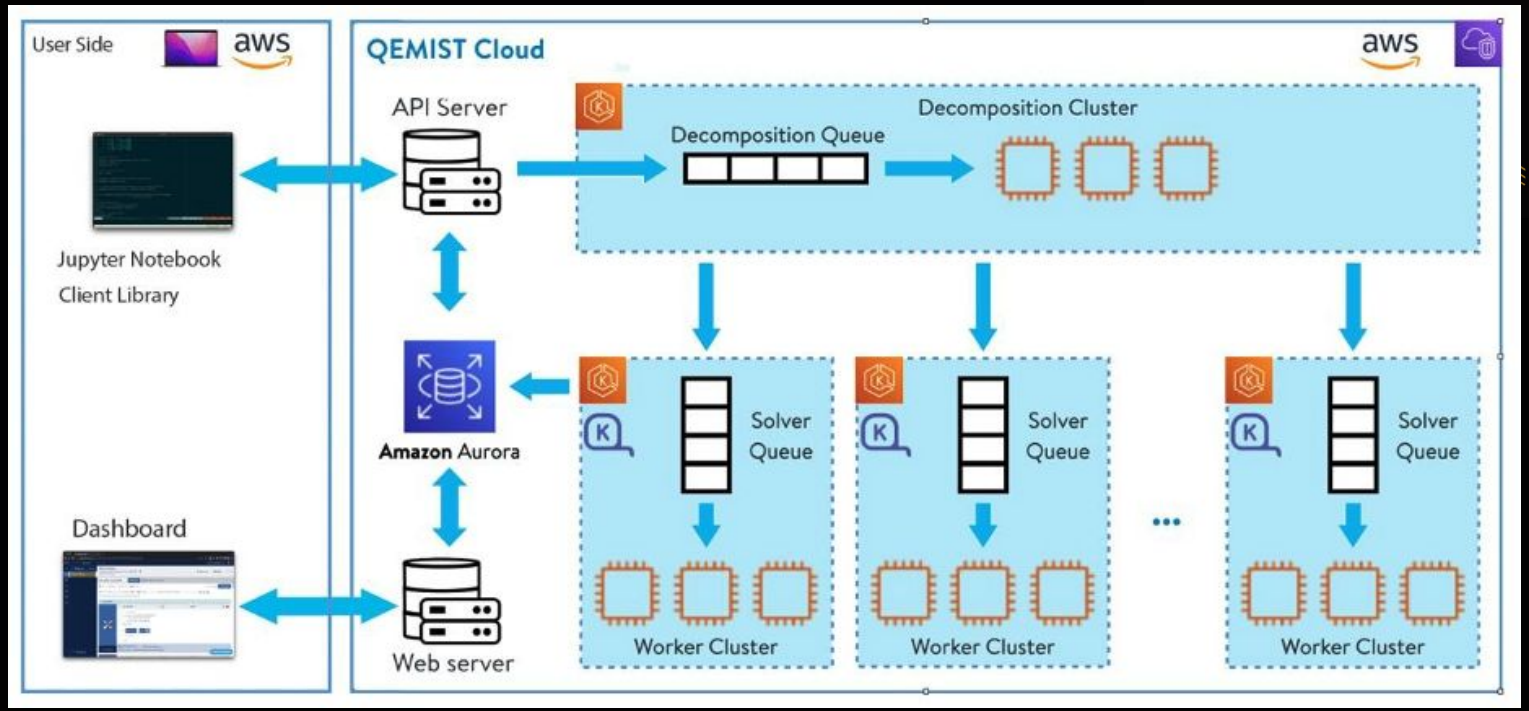
Computational Chemistry 101

- Want to calculate the Potential Energy Surface.A mathematical function that gives the energy of a molecule as a function of its geometry

- The energetics of the C-F bond breaking process are currently being investigated for several PFAS compounds in collaboration with Accenture, AWS and Intel.



General Architecture



5 Journey

Prerequisites

Cost Mitigation
Spot instances, VPC Endpoints

Monitoring
AWS Cloudwatch dashboards, Container logs

Stability
Terraform, CI/CD, Static Code Analysis, Error handling

Logging
AWS CloudWatch Log Insights instead of ELK,

Scalability
Karpenter, KEDA, AWS Aurora Serverless, AWS RDS Proxy

Heartbeat
Race conditions

6 Experiment

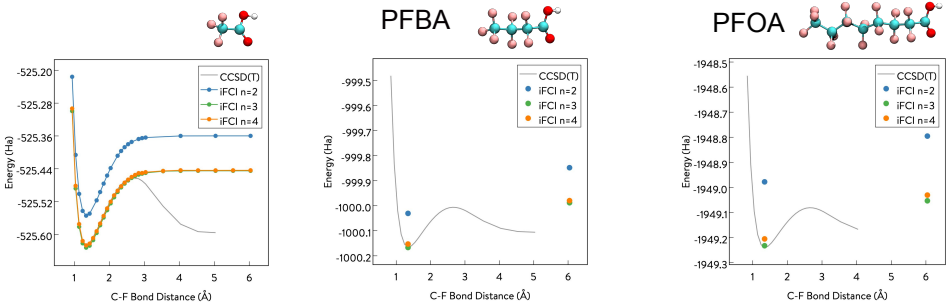
D Day

7 Results

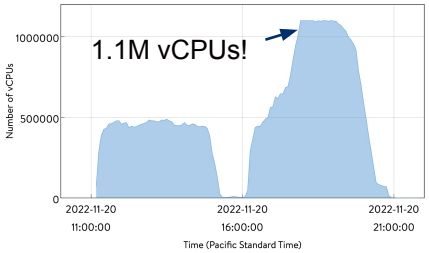
Record-breaking one million core simulation of PFAS chemistry

First quantum chemistry calculation using more than one million vCPUs on the cloud.

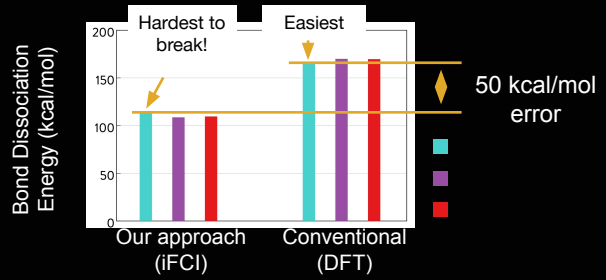
16K vCPUs ⟷ 1.1M vCPUs



Scaling of QEMIST Cloud during the PFOA run



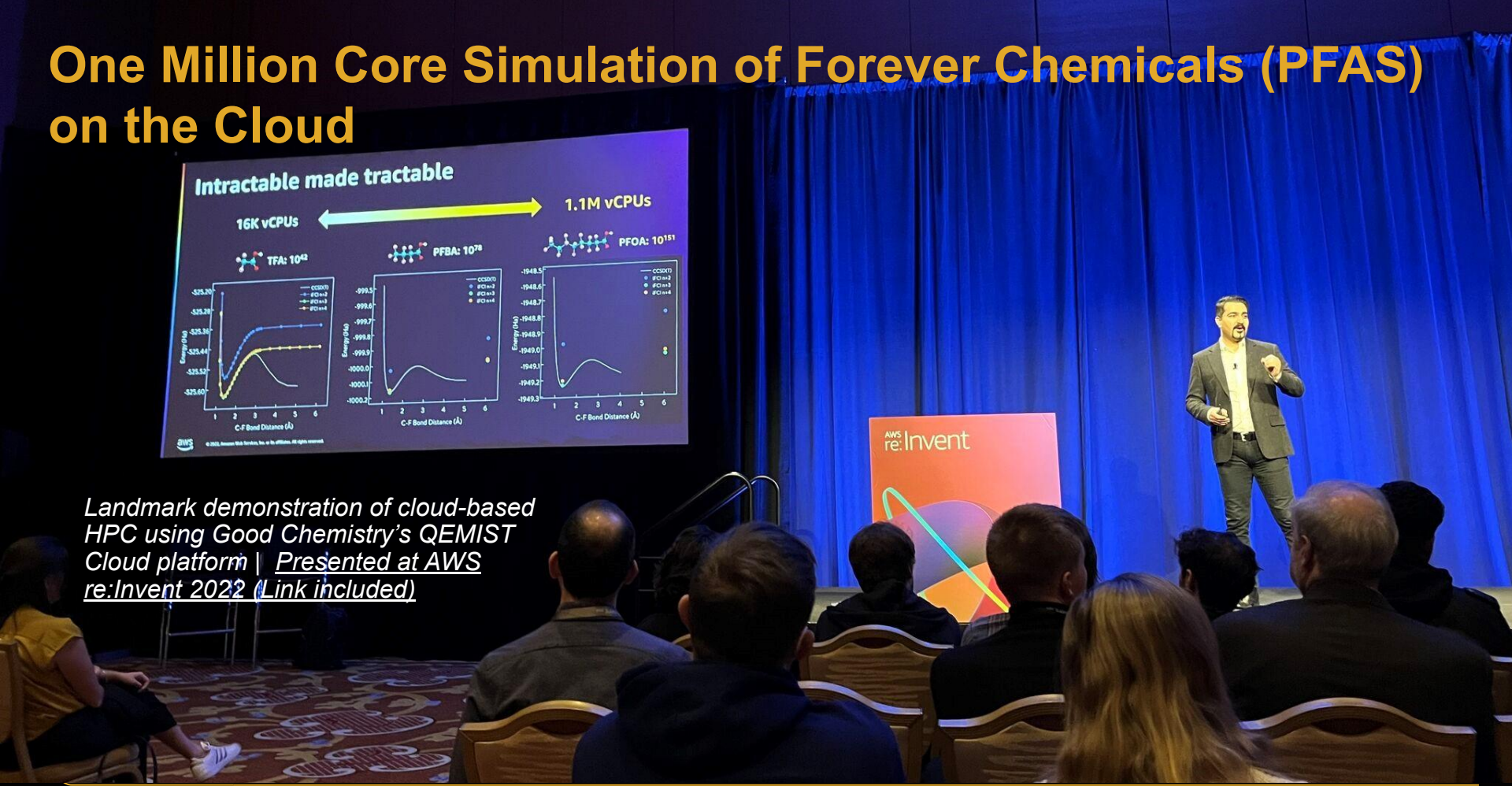
Most accurate quantum chemistry calculation of PFAS bond-breaking energy to date



Uncovered the bond breaking behaviour that differs quantitatively and qualitatively from the conventional methods.

SANDBOXAQ™

One Million Core Simulation of Forever Chemicals (PFAS) on the Cloud

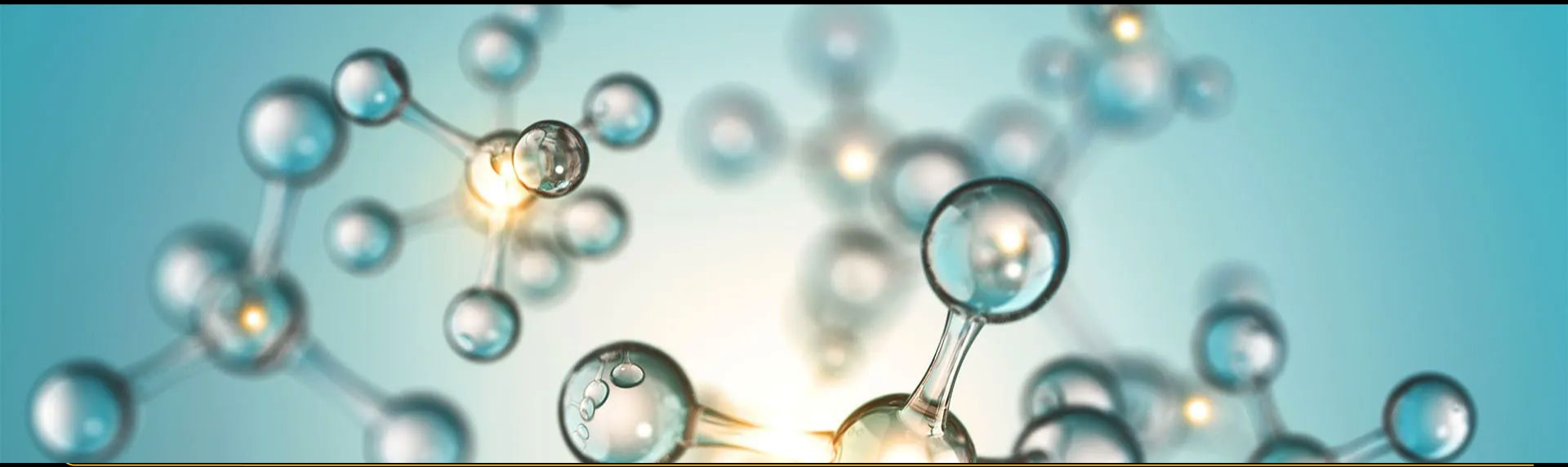


Landmark demonstration of cloud-based HPC using Good Chemistry's QEMIST Cloud platform | Presented at AWS re:Invent 2022 (Link included)

SANDBOXAQ™

A Path to Breaking Up Forever Chemicals

Good Chemistry uses high performance computing in the cloud to accelerate research on removing hazards from the environment



PFAS Related Resources

- Scientific Manuscript
- AWS HPC blogpost
- AWS re:Invent talk
- HPC Wire article
- Wired Article

AWS re:Invent

8 Lessons

Technical Challenges & Lessons Learned

- **Kubernetes:**
  13 clusters is too much!
  Hard to reroute problems
- **Database:**
  Connection limitations
  Race conditions
  Network traffic
- **Data Modelling:**
  Static vs Non Static data
  Nested JSON is the enemy
- **Services:**
  Decoupling needed
- **Error Handling:**
  Duplicate logs made monitoring tougher
  Handling transient errors using retries
- **Memory issues:**
  Problem rerouting
- **Heartbeat:**
  Bottleneck when scaling
  LIFO vs FIFO
- **“Bursty” Traffic:**
  Slow startup time
  KEDA descaling

9 Future

What's Next?

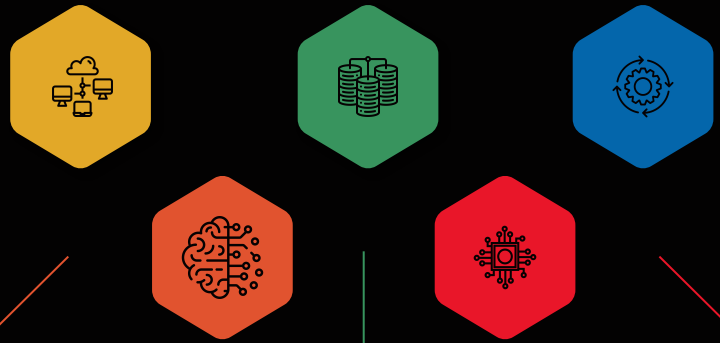
Lots more work to do!

Multi-Cloud

Features, availability, pricing

Serverless

"Bursty" traffic



Polyglot compute

CPU, GPU, QPU

Data Contracts

Enables for a robust microservice architecture

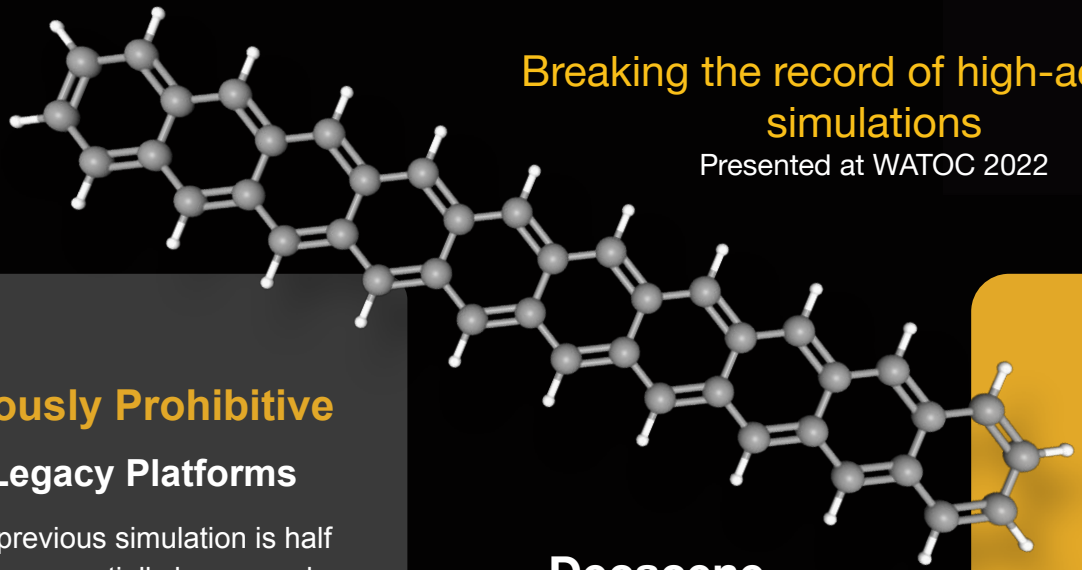
Events

Tell the world I did a thing!

10

Success Stories

Largest molecular simulation with near-exact accuracy



Breaking the record of high-accuracy simulations

Presented at WATOC 2022

Decacene

$C_{42}H_{24}$

Previously Prohibitive

On Legacy Platforms

Largest previous simulation is half the size, exponentially less complex

Made Possible on QEMIST Cloud

Only 11 hours, ~$2K

On commodity hardware, enabled by massive on-demand parallelization

Take your chemistry to any quantum platform

Tangelo

Toolboxes	Algorithms (NISQ and FT)			Problem decomposition
qubit mappings	**Classical** FCI, CCSD…	**Variational** VQE & variants	**Phase Estimation**	ONIOM
post-processing				DMET
measurements	**Projective** imaginary time	**Hamiltonian evolution**	QSP, etc	QM / MM
molecular computation				MI-FNO and iFCI

. . .

Built-in backends		Custom backends
Simulators	**QPUs**	**User-defined**
Qulacs Cirq Qiskit Sympy …	IonQ Braket IBM Q …	You define how computation is carried out.

Amazon Braket

Azure Quantum

IONQ

Cirq

PENNYLANE

IBM Q

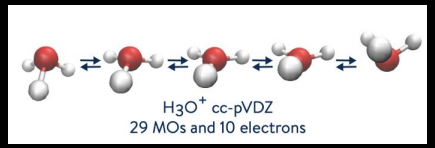
Qiskit

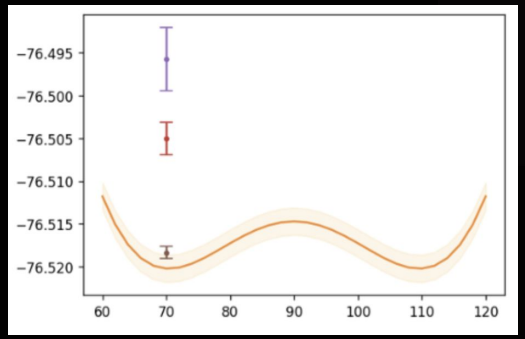
. . .

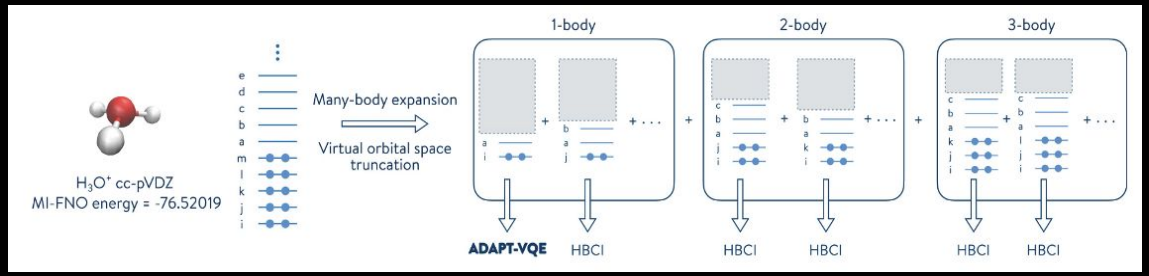
SANDBOXAQ™

Combining QEMIST Cloud, Tangelo and Amazon Braket

- Quantum chemistry with Tangelo and QEMIST Cloud







We are but a small piece of the bigger puzzle…

Additional Resources

- SandboxAQ YouTube Channel

  Dozens of fascinating videos about AQ
- Jack Hidary @ Abundance360

  Our CEO describes in detail how AQ is changing the way we see the world.

SANDBOXAQ™

THANK YOU!

- david.walker@sandboxaq.com
- https://ca.linkedin.com/in/daithi-walker